UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

| APPLICATION NO. | FILING DATE | FIRST NAMED INVENTOR | ATTORNEY DOCKET NO. | CONFIRMATION NO. |
|---|---|---|---|---|
| 09/851,675 | 05/09/2001 | Jan Magnus Stensmo | ARC920000150US1 | 3256 |

7590 09/20/2004

John L. Rogitz
Rogitz & Associates
Suite 3120
750 B Street
San Diego, CA 92101

| EXAMINER | |
|---|---|
| ALBERTALLI, BRIAN LOUIS | |

| ART UNIT | PAPER NUMBER |
|---|---|
| 2655 | |

DATE MAILED: 09/20/2004

Please find below and/or attached an Office communication concerning this application or proceeding.

| Office Action Summary | Application No. | Applicant(s) | |
|---|---|---|---|
| | 09/851,675 | STENSMO, JAN MAGNUS | |
| | Examiner | Art Unit | |
| | Brian L Albertalli | 2655 | |

*-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --*

**Period for Reply**

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

**Status**

1)☒ Responsive to communication(s) filed on *7/26/04*.

2a)☐ This action is **FINAL**. 2b)☒ This action is non-final.

3)☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

**Disposition of Claims**

4)☒ Claim(s) *1-27* is/are pending in the application.

4a) Of the above claim(s) _____ is/are withdrawn from consideration.

5)☐ Claim(s) _____ is/are allowed.

6)☒ Claim(s) *1-27* is/are rejected.

7)☐ Claim(s) _____ is/are objected to.

8)☐ Claim(s) _____ are subject to restriction and/or election requirement.

**Application Papers**

9)☐ The specification is objected to by the Examiner.

10)☐ The drawing(s) filed on _____ is/are: a)☐ accepted or b)☐ objected to by the Examiner.

Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).

Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).

11)☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

**Priority under 35 U.S.C. § 119**

12)☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).

a)☐ All b)☐ Some * c)☐ None of:

1.☐ Certified copies of the priority documents have been received.

2.☐ Certified copies of the priority documents have been received in Application No. _____.

3.☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).

* See the attached detailed Office action for a list of the certified copies not received.

**Attachment(s)**

1) ☒ Notice of References Cited (PTO-892)
2) ☐ Notice of Draftsperson's Patent Drawing Review (PTO-948)
3) ☐ Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) Paper No(s)/Mail Date _____.
4) ☐ Interview Summary (PTO-413) Paper No(s)/Mail Date. _____.
5) ☐ Notice of Informal Patent Application (PTO-152)
6) ☐ Other: _____.

## DETAILED ACTION

### *Response to Arguments*

1. Applicant's arguments with respect to claims 1, 18, 14 and 26 have been considered but are moot in view of the new ground(s) of rejection.

### *Claim Rejections - 35 USC § 103*

2. The text of those sections of Title 35, U.S. Code not included in this action can be found in a prior Office action.

3. Claims 1-4, 6-13, 15, 18-20, and 22-27 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lewis et al. (*A Sequential Algorithm for Training Text Classifiers*), in view of Shavlik et al. (*Intelligent Agents for Web-Based Tasks: An Advice-Taking Approach*).

In regard to claims 1 and 26, Lewis et al. discloses a computer-implemented method (algorithm, Title) for retrieving documents comprising:

inputting the text of one or more documents, wherein each document includes human readable words (section 4.2, the classifier is trained based on words, second paragraph, lines 1-2); and

generating a statistical evaluation of a window wherein the results are not a function of the order of appearance of the words within each window (independence assumption, equation 3, text is categorized according to the components of an observed pattern using Bayesian independence formulation, page 5, equation 6; which is

calculated as being independent of the component pattern order, page 6, section 4.2, first paragraph, equation 7).

Lewis et al. does not disclose creating context windows around each word in each document;

generating an evaluation of the characteristics of all of the windows, wherein the results are not a function of the order of the appearance of words within each window; and

combining the results of the evaluation for each window.

Shavlik et al. discloses a method for retrieving documents comprising:

creating context windows around words in each document (slides a fixed sized window across words in a page, page 3, first column, Scoring Arbitrarily Long Pages with Fixed-Sized Neural Networks section, lines 3-10);

generating an evaluation of the characteristics of all of the windows, wherein the results are not a function of the order of the appearance of words within each window (each window is represented by a "localized bag-of-words" model, wherein word order is lost, page 3, Extracting Features from Web Pages section; page 4, first column, lines 1-5; and Fig. 2).

and combining the results of the evaluation for each window (ScorePage produces a score as the window is slid over the page, page 3, first column, Scoring Arbitrarily Long Pages with Fixed-Sized Neural Networks section, lines 11-14 and Fig. 1).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Lewis et al. to create a context windows around words in each document, generate a statistical evaluation of each window and combine the results of the statistical evaluation of each window, because, as opposed to the standard "bag-of-words" representation of a document, using "localized bag-of-words" representations provides "a richer representation that preserves some word-order information", as taught by Shavlik et al. (page 3, second column, lines 15-17).

Neither Lewis et al. nor Shavlik et al. disclose that a context window is created around *each* word in the document (Lewis et al. only performs a statistical evaluation on the 70% of words most common in a document, while Shavlik et al. discards common stop words).

Official notice is taken that it is well known in the art to utilize *each* word in a document for an evaluation of that document for text retrieval. It would have been obvious to one of ordinary skill in the art at the time of invention to modify the combination of Shavlik et al. and Lewis et al. to include *each* word in the document in a statistical evaluation of that document to ensure that every feature in that document would be included in the statistical evaluation, thereby reducing the chances of retrieving an irrelevant document.

In regard to claim 2 and 27, Lewis et al. discloses determining the likelihood of documents having predetermined characteristics based on the statistical evaluation of a window (each set of words **w** is evaluated to determine which class the set of words

belongs to, the class having been predetermined, page 5, section 4.1, first paragraph; and page 7, section 5.1, last paragraph).

Lewis et al. does not disclose the likelihood is based on the combined statistical evaluation of a plurality of windows, however, as discussed in reference to claim 1, above, the combination of Lewis et al. and Shavlik et al. would determine the likelihood of documents having predetermined characteristics based on the combined statistical evaluation for each window.

In regard to claim 3, Lewis et al. does not disclose assigning a document identifier to each document and context window position; and

determining the document identifier of at least one document having said predetermined characteristics.

Shavlik et al. discloses assigning a document identifier to each document and context window position (URL of page is stored to identify the document and a bag is created for the positions in the sliding window, page 4, first column, line 2; and third paragraph, lines 1-4); and

determining the document identifier of at least one document having said predetermined characteristics (a user inputs relevant information, which is then sent to WAWA, which finds 100 pages rated the highest, page 6, second column, Experiments section, lines 10-17; and page 2, first column, System Description section, second paragraph).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Lewis et al. to assign document identifiers to each document and context window position, and to determine the document identifier of at least one document having predetermined characteristics, so that a document determined to be relevant could be returned to the user and allow the user to access that document with the document identifier.

In regard to claim 4, Lewis et al. discloses defining a plurality of document categories; and determining the category of a particular document based on the statistical evaluation for a window (each set of words **w** is evaluated to determine which class the set of words belongs to, the class having been predetermined, page 5, section 4.1, first paragraph).

Lewis et al. does not disclose the category is determined based on the combined statistical evaluation for a plurality of windows, however, as discussed in reference to claim 1, above, the combination of Lewis et al. and Shavlik et al. would determine the category of a particular document based on the combined statistical evaluation for each window.

In regard to claim 6, Lewis et al. discloses generating a statistical evaluation further includes counting the occurrences (number of instances) of particular words and particular documents and generating totals of the counts (page 6, section 4.2, first paragraph).

In regard to claim 7, Lewis et al. discloses generating counts about singular word occurrences (*d*, the number of features) and about pair-wise occurrences (the number of instances a word is in the positive training set and the number of instances a word is in the negative training set, page 6, section 4.2, first paragraph).

In regard to claim 8, Lewis et al. discloses that the number of distinct words can be very large and that feature selection improves effectiveness (page 6, section 4.2, second paragraph, lines 1-3). Furthermore, Shavlik et al. discloses discarding common words greatly reduces the dimensionality of the classification problem (page 3, second column, second paragraph).

Neither Lewis et al. nor Shavlik et al. specifically discloses pruning the number of pair-wise counts.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the combination of Lewis et al. and Shavlik et al. to prune the number of pair-wise counts in order to reduce the dimensionality of the classification problem and improve the effectiveness of the classification.

In regard to claim 9, neither Lewis et al. nor Shavlik et al. discloses monitoring the amount of memory used for the pair-wise counts and pruning when a predetermined threshold of memory has been exceeded for the pair-wise counts.

However, since both Lewis et al. and Shavlik et al. suggest removing certain words from the evaluation of a document, it would suggest to one of ordinary skill in the art that some features of a document could be disregarded, and the statistical evaluation would still be accurate.

Official notice is taken that it is well known and recognized in the art to monitor the amount of memory used to store information.

It would have been obvious to one of ordinary skill in the art at the time of invention to modify the combination of Lewis et al. and Shavlik et al. to monitor the amount of memory used by the pair-wise counts and prune the pair-wise count when a predetermined threshold of memory had been exceeded, so that a highly accurate statistical evaluation that included all instances of pair-wise counts would occur unless the memory threshold was reached, in which case the number of pair-wise counts would be pruned. Pruning the pair-wise counts would allow the statistical evaluation to continue after the threshold was reached and, as suggested by Lewis et al. and Shavlik et al., would still provide an accurate statistical evaluation.

In regard to claims 10-12, Lewis et al. discloses a method (algorithm, Title) for determining a conditional probability of particular words (**w**) appearing in particular categories based on the counts using a Simple Bayes statistical model (log likelihood ratio from the Bayesian independence formulation, page 5, equation 6). Lewis et al. discloses the method is useful for information retrieval (page 10, lines 1-4).

In regard to claim 13, the combination of Lewis et al. and Shavlik et al., as applied to claim 1, above, discloses in Shavlik et al. that the step of creating context windows around each word further comprises the step of selecting the words appearing before and after each word by a predetermined amount in the document and including those selected words in the window (a fixed size window is slid across the page, and features are defined according to the center of the window, therefore, the window must include words before and after the center word of the window, page 3, first column, coring Arbitrarily Long Pages with Fixed-Sized Neural Networks section, lines 3-10).

In regard to claim 15, the Lewis et al. discloses normalizing the statistical evaluation for the windows (equation 6, parameter *b* serves to dampen extreme log likelihood ratio resulting from independence violations, page 5, section 4.1, last paragraph, lines 1-3). The combination of Lewis et al. and Shavlik et al., as discussed in reference to claim 1, above, would therefore normalize the combined statistical evaluation for the windows.

In regard to claim 18, Lewis et al. discloses:

a means for inputting the text of one or more documents, wherein each document includes human readable words (section 4.2, the classifier is trained based on words, second paragraph, lines 1-2);

means for generating a statistical evaluation of a window wherein the results are not a function of the order of appearance of the words within each window

(independence assumption, equation 3, text is categorized according to the components of an observed pattern using Bayesian independence formulation, page 5, equation 6; which is calculated as being independent of the component pattern order, page 6, section 4.2, first paragraph, equation 7); and

means for determining the likelihood of documents having predetermined characteristics based on the statistical evaluation of a window (each set of words **w** is evaluated to determine which class the set of words belongs to, the class having been predetermined, page 5, section 4.1, first paragraph; and page 7, section 5.1, last paragraph).

Lewis et al. does not disclose the category is determined based on the combined statistical evaluation for a plurality of windows, however, as discussed in reference to claim 1, above, the combination of Lewis et al. and Shavlik et al. would determine the category of a particular document based on the combined statistical evaluation for each window.

Lewis et al. also does not disclose a storage unit for receiving a plurality of documents;

means for creating context windows around each word in each document;

means for generating an evaluation of the characteristics of all of the windows, wherein the results are not a function of the order of the appearance of words within each window; and

means for combining the results of the evaluation for each window.

Shavlik et al. discloses a method for retrieving documents comprising:

means for creating context windows around words in each document (slides a fixed sized window across words in a page, page 3, first column, Scoring Arbitrarily Long Pages with Fixed-Sized Neural Networks section, lines 3-10);

means for generating an evaluation of the characteristics of all of the windows, wherein the results are not a function of the order of the appearance of words within each window (each window is represented by a "localized bag-of-words" model, wherein word order is lost, page 3, Extracting Features from Web Pages section; page 4, first column, lines 1-5; and Fig. 2);

means for combining the results of the evaluation for each window (ScorePage produces a score as the window is slid over the page, page 3, first column, Scoring Arbitrarily Long Pages with Fixed-Sized Neural Networks section, lines 11-14 and Fig. 1);

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Lewis et al. to include means for creating context windows around words in each document, means for generating a statistical evaluation of each window and combine the results of the statistical evaluation of each window, because, as opposed to the standard "bag-of-words" representation of a document, using "localized bag-of-words" representations provides "a richer representation that preserves some word-order information", as taught by Shavlik et al. (page 3, second column, lines 15-17).

Neither Lewis et al. nor Shavlik et al. disclose that a context window is created around *each* word in the document (Lewis et al. only performs a statistical evaluation on

the 70% of words most common in a document, while Shavlik et al. discards common stop words).

Official notice is taken that it is well known in the art to utilize *each* word in a document for an evaluation of that document for text retrieval. It would have been obvious to one of ordinary skill in the art at the time of invention to modify the combination of Shavlik et al. and Lewis et al. to include *each* word in the document in a statistical evaluation of that document to ensure that every feature in that document would be included in the statistical evaluation, thereby reducing the chances of retrieving an irrelevant document.

Furthermore, official notice is taken that it is well known in the art to include a storage unit in a computer system to store documents, therefore it would have been obvious to one of ordinary skill in the art at the time of invention to modify Lewis et al. to include a storage unit for receiving and storing a plurality of documents.

In regard to claim 19, Lewis et al. does not disclose a document identifier assigned to each document and context window position; and

means for determining the document identifier of at least one document having said predetermined characteristics.

Shavlik et al. discloses a document identifier assigned to each document and context window position (URL of page is stored to identify the document and a bag is created for the positions in the sliding window, page 4, first column, line 2; and third paragraph, lines 1-4); and

means for determining the document identifier of at least one document having said predetermined characteristics (a user inputs relevant information, which is then sent to WAWA, which finds 100 pages rated the highest, page 6, second column, Experiments section, lines 10-17; and page 2, first column, System Description section, second paragraph).

It would have been obvious to one of ordinary skill in the art at the time of invention to modify Lewis et al. to assign document identifiers to each document and context window position, and to include means for determining the document identifier of at least one document having predetermined characteristics, so that a document determined to be relevant could be returned to the user and allow the user to access that document with the document identifier.

In regard to claim 20, Lewis et al. discloses means for defining a plurality of document categories; and determining the category of a particular document based on the statistical evaluation for a window (each set of words **w** is evaluated to determine which class the set of words belongs to, the class having been predetermined, page 5, section 4.1, first paragraph).

Lewis et al. does not disclose the category is determined based on the combined statistical evaluation for a plurality of windows, however, as discussed in reference to claim 18, above, the combination of Lewis et al. and Shavlik et al. would include means for determining the category of a particular document based on the combined statistical evaluation for each window.

In regard to claim 22, Lewis et al. discloses generating a statistical evaluation further includes counting the occurrences (number of instances) of particular words and particular documents and generating totals of the counts (page 6, section 4.2, first paragraph).

In regard to claims 23 and 24, Lewis et al. discloses a means (algorithm, Title) for determining a conditional probability of particular words (**w**) appearing in particular categories based on the counts using a Simple Bayes statistical model (log likelihood ratio from the Bayesian independence formulation, page 5, equation 6). Lewis et al. discloses the method is useful for information retrieval (page 10, lines 1-4)

In regard to claim 25, the combination of Lewis et al. and Shavlik et al., as applied to claim 1, above, discloses in Shavlik et al. that the means for creating context windows around each word further comprises the step of selecting the words appearing before and after each word by a predetermined amount in the document and including those selected words in the window (a fixed size window is slid across the page, and features are defined according to the center of the window, therefore, the window must include words before and after the center word of the window, page 3, first column, coring Arbitrarily Long Pages with Fixed-Sized Neural Networks section, lines 3-10).

4. Claims 5, 14, and 21 are rejected under 35 U.S.C. 103(a) as being unpatentable over Lewis et al., in view of Shavlik et al., and further in view of Pedersen (*A Simple Approach to Building Ensembles of Naïve Bayesian Classifiers for Word Sense Disambiguation*).

In regard to claim 5, neither Lewis et al. nor Shavlik et al. discloses determining the word that is in the center of a particular window based on the combined statistical evaluation for each window.

Pedersen discloses a method of word sense disambiguation that creates unordered windows around a center word and determines the word that is in the center of the window based on a statistical evaluation of each window (page 2, section 2.1, third paragraph). A word sense can be used to determine what word in the center of a context.

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Lewis et al. and Shavlik et al. to determine the word that was in the center of the window since correct word sense disambiguation increases the performance of machine translation systems as well as information retrieval systems.

In regard to claim 14, neither Lewis et al. nor Shavlik et al. disclose the word around which each window is created is not included in the window.

Pederson discloses the word around which each window is created is not included in the window (windows are to the left and right of an ambiguous word, page 2, section 2.1, third paragraph).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Lewis et al. and Shavlik et al. to not include the center word in each window, since, as discussed in reference to claim 5, above, it is possible to determine what the center word is without including it in the window. Therefore, the center word would not have to be evaluated, thereby saving memory space and computation time.

In regard to claim 21, neither Lewis et al. nor Shavlik et al. discloses means for determining the word that is in the center of a particular window based on the combined statistical evaluation for each window.

Pedersen discloses means for word sense disambiguation that creates unordered windows around a center word and determines the word that is in the center of the window based on a statistical evaluation of each window (page 2, section 2.1, third paragraph). A word sense can be used to determine what word in the center of a context.

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Lewis et al. and Shavlik et al. to include means for determining the word that was in the center of the window since correct word

sense disambiguation increases the performance of machine translation systems as well as information retrieval systems.

5. Claim 16 is rejected under 35 U.S.C. 103(a) as being unpatentable over Lewis et al., in view of Shavlik et al., and further in view of Dumais et al (U.S. Patent 6,192,360).

Neither Lewis et al. nor Shavlik et al. disclose determining a measure of mutual information.

Dumais et al. discloses evaluating a given feature in a textual document by determining a measure of mutual information (column 12, lines 59-67 and column 13, lines 1-54).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Lewis et al. and Shavlik et al., so that a measure of mutual information was determined in the statistical evaluation of a window, so that windows (features) with the highest mutual information values could be kept, while other windows (features) were not considered, as taught by Dumais et al. (column 13, lines 47-54), thereby reducing the amount of memory needed for the statistical evaluation.

6. Claim 17 is rejected under 35 U.S.C. 103(a) as being unpatentable over Lewis et al., in view of Shavlik et al., and further in view of Caid et al. (U.S. Patent 5,619,709).

Neither Lewis et al. nor Shavlik et al. discloses the step of combining includes averaging probability assessments.

Caid et al. discloses the step of combining includes averaging probability assessments (context vectors are combined by a weighted sum, fig. 4, 401 and 402, column 9, lines 11-15).

It would have been obvious to one of ordinary skill in the art at the time of invention to further modify the combination of Lewis et al. and Shavlik et al. so the step of combining averaged the probability assessments, to create a good representation of the document in it's entirety, without giving too much weight to sections of the text in which the same word was used several times without being used again in the rest of the text.

***Conclusion***

7. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Brewster et al. (U.S. Patent 6,070,133) discloses a method that slides an overlapping window across a text and generates frequency counts of the words in the window without weighting the position of the words within the window. Moreno et al. (U.S. Patent 6,772,120) discloses a method for segmenting text streams that slides a window across the text that uses Naive Bayes to evaluate the windows. Shavlik et al. (*An Instructable, Adaptive Interface for Discovering and Monitoring Information on the World-Wide Web*) discloses an additional description of the Wisconsin Adaptive Web Assistant.

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Brian L Albertalli whose telephone number is (703) 305-1817. The examiner can normally be reached on Monday - Friday, 8:30 AM - 5:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Talivaldis Smits can be reached on (703) 305-3011. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see http://pair-direct.uspto.gov. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).

BLA 9/8/04

TALIVALDIS IVARS SMITS
PRIMARY EXAMINER